(51) International Patent Classification 5 :		(11) International Publication Number:	**WO 90/08360**
G06F 15/403	**A1**	(43) International Publication Date:	26 July 1990 (26.07.90)

(21) International Application Number: PCT/US90/00037

(22) International Filing Date: 9 January 1990 (09.01.90)

(30) Priority data:
296,146   12 January 1989 (12.01.89)   US

(71) Applicant: TELEBASE SYSTEMS, INC. [US/US]; 763 W. Lancaster Ave., Bryn Mawr, PA 19010 (US).

(72) Inventors: MEYER, Daniel, E. ; 276 Upper Gulph Road, Radnor, PA 19087 (US). KOLLIN, Richard, P. ; 1239 Remington Road, Wynnewood, PA 19096 (US). FRANCIS, Gerald, A. ; 108 Llanfair Road, Bala Cynwyd, PA 19004 (US).

(74) Agent: EILBERG, William, H.; 820 Homestead Road, Jenkintown, PA 19046 (US).

(81) Designated States: AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CA, CF (OAPI patent), CG (OAPI patent), CH (European patent), CM (OAPI patent), DE (European patent), DK, DK (European patent), ES (European patent), FI, FR (European patent), GA (OAPI patent), GB (European patent), HU, IT (European patent), JP, KP, KR, LK, LU (European patent), MC, MG, ML (OAPI patent), MR (OAPI patent), MW, NL (European patent), NO, RO, SD, SE (European patent), SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent).

Published
With international search report.

(54) Title: SYSTEM AND METHOD FOR RETRIEVING INFORMATION FROM A PLURALITY OF DATABASES

(57) Abstract

This invention enables a user (5) to obtain information from a large number of commercial databases (20-24). In practicing the invention, the user (5) selects an area of interest (40) and enters a search request (42). The search request (42) includes at least one word for which the user (5) desires to search. In one embodiment of the invention, the system selects a set of at least two databases (two of 20-24), automatically executes the search request (46) in each database, and presents the results to the user (5, 48). In another aspect of the invention, the user (5) selects a database directly, and employs a set of standardized commands (see figure 4) for any database selected. The system translates these standardized commands into the equivalent commands recognized by each database, without the intervention or knowledge of the user (5). The user (5) can thus communicate with a variety of databases (20-24) using the same command set (see figure 4). In another embodiment, the invention guides the user (4) in reformulating a search which retrieved either no documents or too many documents. The invention also includes a method of determining which of the retrieved documents are likely to be the most relevant.
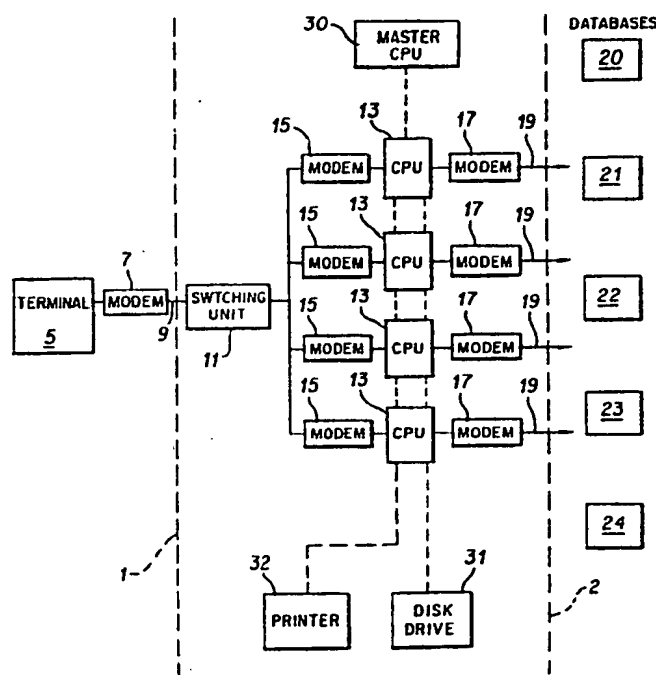
SYSTEM AND METHOD FOR RETRIEVING INFORMATION
FROM A PLURALITY OF DATABASES

CROSS-REFERENCE TO PRIOR APPLICATIONS

This is a continuation-in-part of U.S. Patent Application Serial No. 231,055, filed August 11, 1988, entitled "System for Retrieving Information From a Plurality of Remote Databases", which is a continuation of Serial No. 664,167, filed Oct. 24, 1984, now U.S. Patent No. 4,774,655.

BACKGROUND OF THE INVENTION

This invention relates to the field of information retrieval, and especially the retrieval of information from one or more databases.

The field of information retrieval has advanced to the point where anyone having a personal computer, a modem, and access to a telephone line can obtain information on virtually any topic, from thousands of commercial databases, without leaving home. Many owners of large compilations of information have provided their information in the form of computer databases, and these databases can be interrogated, for a fee, by a remote user.

Groups of related and unrelated databases are sometimes arranged in "families". A single vendor, or "host", may provide access to a family of databases, all of which can be reached through the same telephone number, and with the same identification number and password. For example, Dialog Information Services, Inc., of California, provides access to a large family of databases, dealing with many different subjects, under the service mark DIALOG. In this specification, the term "family" means a set of related or unrelated databases available on one particular host. As used herein, the terms "family" and "host" are virtually interchangeable.

command set, and a different field structure, since each database is, in general, created by a different entity. Thus, to use all of the hundreds or thousands of available databases, it would be necessary to learn all of the rules for syntax, command sets, and field structures for each one.

The problem of nonuniformity of database languages was ameliorated somewhat by the creation of database families, such as DIALOG, mentioned above. Within a particular family, the syntax and command sets are generally the same for each database. As long as a user wishes to search only within one database family, the user needs to know only one search syntax and one command set. Of course, the user would also need a database catalog, in order to know what databases are available.

However, even within a given database family, the field structures of the various databases are, in general, very different. Moreover, if one desires to search in other families of databases, or in single databases not belonging to a family, one must, in general, learn a new search syntax, a new command set, and a new field structure. It is also necessary to establish a separate account with each different database or database family. And, of course, it is necessary for the user to know, in advance, what databases are available and what they contain.

Database searching presents additional problems not discussed above. Frequently, a search yields no documents, or "postings". The null result could be due to misspelled words in the search request, or to other factors. Just as frustrating is the case where a search produces too many postings. It is tedious and expensive to browse through one hundred or more documents retrieved from a database search. Even if the number of postings is not unreasonably large, it is still necessary to browse

The present specification discloses enhancements to the basic system described in the cited patent. One of these enhancements allows a user to search through a large number of databases in one session. Another embodiment improves on the standardization described in the cited patent, by enabling the user to search through many different selected databases using a standardized set of commands. Other features reduce the likelihood of null search results, and also deal with the problem of too many retrieved documents. Another feature provides an economical method of determining, in advance, which retrieved documents are likely to be the most relevant. With these features, the system is even more convenient to use than systems of the prior art, and database searches are more likely to yield meaningful results.

SUMMARY OF THE INVENTION

The present invention enables a user to obtain information from one or more databases. The user communicates with a central computer, programmed according to the invention, and the computer establishes direct communication with the databases. In general, the user is not located at the same place as the central computer, and the central computer is not at the same location as the various databases. However, the invention is not limited by the distances between the user, the central computer, and the databases.

In one aspect of the invention, the user first transmits an area of interest and a search request to the central computer. The area of interest can be selected from a menu of available subject areas. The search request includes one or more words, connected by logical operators, such as Boolean operators or proximity connectors. The object of

selected by the user. Thus, although the user may want to select a particular database for searching, the user need not learn the particular command set for that database. The central computer establishes the connection with the selected database, and the user does not directly deal with that database.

The latter embodiment may be combined with other embodiments. Thus, the standardized command set can be used even in the case where the system chooses the database for the user. Also, the user may select a database but leave the translation of both the search request and the commands to the system. In the latter case, the user need not know anything about the database other than its name.

In another embodiment of the invention, the system guides the user in reformulating a search, in those cases where a search produces no postings or too many postings. The system identifies words of the search request which were not found in any document, and gives the user the opportunity to modify or delete those words. In cases where each word of the search request yielded postings, but the overall request did not yield any postings, the system provides suggestions to the user for broadening the search. In the case of too many postings, the system provides the user with suggestions on how to narrow the search, such as by imposing more stringent field or proximity restrictions on portions of the search request.

The invention also includes a system and method for determining which retrieved documents are likely to be the most relevant. This method is performed without browsing through the actual documents. Instead, the system considers, for example, the fields in which the search terms were found. Thus, if a search term is found in the title of the document, it is likely that the document is more relevant than if the

It is another object to increase the ease and effectiveness of database searching.

Other objects and advantages of the invention will be apparent to those skilled in the art from a reading of the following brief description of the drawings, the detailed description of the invention, and the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a block diagram showing a possible configuration of the system of the present invention.

Figure 2 is a flow chart showing the operation of the embodiment of the invention wherein the system executes a search automatically in each of several databases.

Figure 3 is a block diagram showing an alternative configuration of the system, wherein the central computer can perform searches simultaneously in more than one database.

Figure 4 is a chart containing a hypothetical standardized set of commands, used for searching in databases, and showing the meaning of each command.

Figure 5 is a flow chart illustrating an embodiment of the invention which assists the user in reformulating a search that has retrieved too few or too many documents.

which are available, such as packet assemblers and disassemblers ("pads"). Also, the connection between terminal 5 and computer 13 may include a conventional analog connection between the terminal and a given node in a telecommunications network, and a digital connection between that node and one of computers 13. Such alternative arrangements are within the scope of the present invention.

Each of the databases 20-24 represents, in general, a separate and independent computer system at some remote location. The information stored in each database may be kept on disks or other storage media, and the searching through the respective databases is governed by the particular computer system for that database. The system of the present invention assumes the existence of such multiple databases, and operates with any or all of them, regardless of the specific configuration of the computer system of each database. The system communicates with these databases in the same manner as would any other user of such databases.

Connected in parallel to all the computers 13 is master CPU 30. Master CPU 30 is a computer which controls the overall operation of the system. The master CPU continually checks that all computers 13 are operating properly. It also can be programmed, in conjunction with switching unit 11, to guide an incoming user call to the appropriate computer. The master CPU 30 also serves to direct the various computers to retrieve information from disk 31, or to print billing information on printer 32.

The components shown between dotted lines 1 and 2 indicate the components which are "local", i.e. grouped at a central location. The components shown outside the dotted lines are "remote", i.e. located elsewhere. Normally, terminal 5 will be located at a substantial distance from the system, possibly thousands of miles away. Databases 20-24 are

13

The system described in the cited patent then automatically dials the telephone number of the selected database, and establishes a connection, using an identification number and password which has been stored in memory. Note that it is the system, not the ultimate user, which is the customer of the database. The system automatically translates the search request into the search syntax of the selected database, and transmits the search request to the database. The system may then download, to its own memory, some or all of the results of the search, and terminates the connection with the database. The system then displays the downloaded results of the search to the user. This displaying step may include allowing the user to browse electronically through one or more articles or other documents retrieved in the search. This browsing is therefore done after the system is disconnected ("off-line") from the database. The system can also print a bill, using printer 32, based on credit card information previously supplied by the user. Note that the user need establish only one account, i.e. a credit card account, to gain access to a wide variety of databases. The individual databases do not "see" the user as a customer.

Further variations to the above-described process are explained in the cited patent. In general, each such variation uses substantially the same general arrangement of hardware, the difference being in the programming of the computers.

Some databases are included in more than one database family, i.e. they are present on more than one host. The database selection step may therefore also include the step of choosing a host on which to search.

The choice of a host can be made in the following manner. For each database, the system stores a list of all hosts on which that database is available. The hosts on each list are ranked in a predetermined order

possible to obtain further access to the host for this reason.   Of

course, the user is unaware of which host the system chooses.

Figure 2 is a flow chart illustrating the operation of an embodiment

of the present invention which is a modification of the procedures de-

scribed above and in the cited patent.   It is assumed, in Figure 2, that

the user has already established connection with the system, and has

transmitted satisfactory credit card information, or other identifica-

tion.   Then, in block 40, the user begins by selecting an area of inter-

est from a menu displayed by the system.   In block 42, the user enters a

search request.   The functions of blocks 40 and 42 could be performed in

reverse order, if desired.   In block 44, the system chooses a set of

databases in which searches will be made.   The system stores, in its mem-

ory, a data file comprising a list of areas of interest and databases.

With each area of interest, the system associates a set of databases, all

of which contain information relating to that area.   This set of databas-

es is fixed for each possible selection of area of interest by the user.

Thus, as in the cited patent, the user need not know, in advance, what

databases to search, and the selection of databases is entirely auto-

matic.

In block 46, the system then searches through each database in the

selected set.   Each search may be performed in substantially the same

manner as described for an individual search in the cited patent.   That

is, the system automatically translates the search request to conform to

the syntax of the database being searched, establishes connection with

the database (transmitting an appropriate identification number and pass-

word, if necessary), and transmits the translated search request to the

database.   If a particular database is a member of a family, the system

gains access to the family, choosing that family according to the method

loads the selected items to its memory, as indicated in block 60, and terminates the connection with the database. In block 62, the system displays the retrieved documents to the user. The function of block 62 may include an interactive display, allowing the user to browse electronically through the document(s). During this displaying step, the user remains connected to the system but the system is not connected to any database, so no additional database charges are incurred. When the user is finished viewing the item, the program can return to test 50, and the document-viewing process can be repeated.

In the procedure discussed above, the system searches a set of databases one at a time. It is also possible to search through many databases virtually simultaneously. However, in the latter case, it would be necessary for each computers 13 to be connected, through separate modems, to a plurality of databases. Alternatively, computers 13 can be arranged as one larger computer. In either case, the computer would be subdivided into separate processors, or otherwise programmed on a time-sharing basis, so that signals could be passed back and forth between the system and each of several databases virtually simultaneously. The computer could also remain connected to all of the databases continuously, especially if there is a constant stream of search requests covering all or most of the available databases or database families.

Figure 3 illustrates one possible arrangement for the latter embodiment. Terminal 70 is connected, by a telephone line, to CPU 72, which can establish connections with databases such as 78, 80, and 82. Dotted lines 74 and 76 indicate the boundaries between the system and the components external to the system. CPU 72 can send data through any of modems 84, 86, 88, 90, and 92. The CPU automatically connects to as many modems as it needs to do the simultaneous searches. The remaining modems can be

level of standardization for database searching. In this embodiment, the user employs one standardized set of commands for conducting searches in disparate databases. Before describing this embodiment, it is helpful to explain the concept of a "command set" for database searching.

As explained above, each commercial database, or family of databases, has its own rules of syntax which governs the construction of search requests. Each set of rules of syntax includes a set of logical operators used to connect the terms of a search request. Examples of rules of syntax for various databases or database families are given in U.S. Patent No. 4,774,655.

It is not enough to learn the search syntax for a particular database or database family. It is also necessary to know a set of commands needed for conducting a search. Such commands are used for such functions as instructing the system to begin a search, displaying the results of a search, providing information on the types of searching available, terminating a research session, and more. In general, each database, or database family, uses a different set of commands.

In the present invention, the user transmits commands selected from a standardized set. The system translates the search request into the equivalent command applicable to the database or database family being searched.

Figure 4 is a table showing a hypothetical standardized set of commands, with a brief indication of the meaning of each. The table also shows the equivalents of each command in the BRS, DIALOG, and VU/TEXT families of databases. For example, the hypothetical standardized command which directs the system to search the database is "FIND". In DIALOG, the corresponding command is "SELECT". In BRS, it is "..SEARCH". In the hypothetical standardized command set, "SHOW" is used to display

mands is in the embodiment of the cited patent wherein the user selects a particular database. The user may know the names and general coverage of a variety of databases, but may not know (and may not want to learn) the command sets appropriate for each database. The system of the present invention can therefore translate standardized commands into the commands recognized by the selected database or database family. The user conducts the search while the system remains connected, or "on-line", with respect to the database. From the viewpoint of the user, it is as if he or she is directly connected to that database. However, the user employs the standardized command set, while the system acts as an intermediary, translating all commands into the commands appropriate to the particular database.

Note that some of the standardized commands (such as "FIND", in the example given above), relate to the process of performing the search, whereas other commands (such as "SHOW", in the above example) relate to the displaying of results. Thus, not all of the standardized commands would be available to the user, in every variation of the invention. For example, if the system selects databases and performs searches entirely automatically, the user would never need to enter a "FIND" command, as the system would do this automatically. But the standardized command set could still be used for directing the display of search results, after the searching is concluded. After the system disconnects from the database and displays the number of documents retrieved, the user can then transmit, for example, a "SHOW" command, followed by the number of the document desired to be displayed. The system would then reconnect to the database, perform the search again, download the requested document, disconnect from the database, and display the document to the user. The above-described arrangement could be used with the embodiment of Figure

format (such as "AU=SMITH") and using a standardized set of field tags. The expression "AU=SMITH" is called a "field command" in this specification. The system can then translate the command into the format appropriate to the database being searched. Also, the system can be programmed to translate the operand (e.g. "SMITH") into an format appropriate to a particular database. For example, some databases contain author information with the last name first, and others place the last name last. The system could be programmed to "know" that, for certain databases, the standardized command "AU=JOHN SMITH" must be translated into "AU=SMITH, JOHN".

The invention thus includes at least three possible levels of translation. The system can translate search requests, commands, and field commands. Any combination of these types of translations can be incorporated into a given embodiment of the invention. Thus, for example, one can provide a system in which the user, not the system, selects a database, wherein the user is expected to know the search syntax of that database, but wherein the system translates search commands from a standardized set. This variation is appropriate for sophisticated users who are familiar with the search syntax of their favorite databases, but who do not want to memorize different sets of search commands. One can even provide standardization of field commands only. The invention is not limited to the above-described combinations, however.

One of the problems in database searching is the retrieval of too few or too many documents, or "postings". As used herein, a "posting" occurs when a word in the search request is found in a document contained in the database. The term "document" is used herein in a general sense, and includes a record retrievable from a database, whether it be an article, a patent, or information recorded in any other format. A search

number of words of each other), and if, say, the third word is not in the database, the search will yield no postings because of that third word. In the latter example, the system will know that it was the third word which caused the search to fail. The system then displays a message to the user, stating that there were no postings for this word. The system asks the user if the word is spelled correctly or whether it was entered in the proper format. By "format" it is meant that the word may have been restricted to a particular field (e.g. title, author, or abstract), and the field designator may be incorrectly entered. If the spelling or format is incorrect, the user can enter the corrected term, and the search will be executed again.

If the spelling and format are already correct, the system gives the user suggestions for modifying the search. These suggestions can include 1) entering a related term instead of the term that caused the search to fail, 2) relaxing field restrictions (e.g. searching for all occurrences of the word instead of limiting the search to the title, author, or abstract), and 3) deleting the term from the search. The user may also choose to abandon the search at this point. The suggestions are preferably arranged in a menu, and the user can easily make one or more choices.

If there are more than one terms that would cause the search to fail, the system would attempt to resolve each problem separately before the search is resubmitted. That is, the system performs all the operations described above for each search term or unit which yielded no postings, before the search can be performed again.

The second category of searches which yield no postings is called "Condition B". In Condition B, all of the units of the search request yield postings, but the full search request yields no postings. In many

27

replace the search with "LEVERAGED AND BUYOUT", which will retrieve all documents containing both words in any location within the document. Finally, the user is also given the opportunity to abandon the search entirely.

The above examples of search syntax are only hypothetical. In practice, the system could use any other means of informing the user about the restrictions that were initially and automatically placed on the search, and can then give the user the chance to relax such restrictions.

Note that in the cases of both Condition A and Condition B, the system identifies the probable points at which the search failed, and offers options for correction.

The system is programmed to display to the user only the options appropriate to a particular search. Thus, for example, if no field restrictions were entered by the system, the user would not be given the option of broadening the field restrictions. If there are no proximity connectors in the search, the user would not be asked to broaden them. Also, if the search request consists of only one term, the option to delete a term would not be presented.

It is quite possible that the user could resolve a Condition A problem and then encounter Condition B. Also, resolution of either Condition A or Condition B could result in the problem of too many postings, to be discussed below. In anticipation of either case, the system should be programmed to place a limit on the number of failed searches that can be performed by one user in one session.

The problem of too many postings is almost as serious as the problem of no postings at all. When the number of postings is greater than 99 (or any other arbitrarily selected number), the system presents a menu of the following choices to the user. First, the user can choose to view

user may abandon the search, in block 142, or enter a modification, in block 144.

If the search yields a nonzero number of postings, the system checks whether there are too many postings, in test 146. If not, the system displays the results, in block 148. Block 148 can include any other display steps that may be desired. If there are too many postings, then the system displays the number of postings in block 150 and asks the user for a choice, in block 152. The user may abandon the search, through test 154 and block 156, or may modify the search in block 158.

Another embodiment of the invention is useful both in the case of too many postings and in the case of a "successful" search. In this embodiment, the system ranks the retrieved documents in order of apparent relevance. The ranking is done without actually browsing through the documents.

The principle used in ranking retrieved documents can be illustrated with a simple example. Suppose the user wants information on laptop computers. The system searches for documents containing "LAPTOP" and "COMPUTERS" within two words of each other. Suppose that 87 documents are retrieved. Now suppose that the search is narrowed by specifying that both words must appear in the title. Now, the number of documents may be reduced to 23. Suppose further that the search is again narrowed by specifying that the words be within one word of each other. Now, the number of retrieved documents is reduced to 12. It is very likely that these 12 documents are the most relevant of the original 87.

In general, it is apparent that a document that contains a word of the search request in its title is probably more relevant than a document containing the word in the abstract or elsewhere. A document in which two words of a search request are adjacent is probably more relevant than

tice, that only a relatively small number of indicators of relevance are necessary. The most useful criterion is to determine the field in which the search term appears. The number of terms appearing in a given field has also been found to be a useful criterion of relevance.

It is important to note that the ranking of documents by relevance is done without actually viewing the documents. On the contrary, the ranking is done only by determining which terms, or combinations of terms, appear in which portions of the documents. Such information can almost always be obtained without examining the document. Even if the database does not display intermediate steps, one can "probe" the database by entering a search for a given word, restricted to a given field, and observing the number of documents retrieved.

In general, in order to rank the documents by relevance, it may be necessary for the system to perform a search, for each given word, more than once. Thus, the system might search for occurrences of a given word, and then might repeat the same search, limiting the second search to, say, the "title" field. While the use of such duplicate searching adds somewhat to the cost of the search, the cost of search time is usually relatively small compared to the cost of displaying search results. Thus, for a slightly higher cost of doing the search, one can obtain considerably more information about the relevance of the documents retrieved, and can thereby avoid the more significant costs of displaying documents that turn out to be nonrelevant.

A method of ranking retrieved documents according to assumed relevance can be summarized as follows. Suppose that the searcher wants to find information about parallel processing computers. The system performs a search for documents containing the words "PARALLEL", "PROCESSING", and "COMPUTERS". It also performs searches for documents which

33

7. Automatic generation of field commands to be used
    in conducting a search

8. Translation of standardized field commands into
    format of database or database family

9. Automatic substitution of alternative form of
    field commands

10. Automatic generation of suggestions for reformulation
    of search requests in cases where no documents or too
    many documents were retrieved

11. Automatic generation of list of retrieved documents
    in the order of their presumed relevance

To the above list, one can add alternative arrangements for format-
ting and displaying the results of searches. Moreover, one can program
the system to repeat the same search periodically to determine whether
relevant documents may have been added to the databases under review.

The features of the above list can be present in virtually all pos-
sible combinations. All of these features have been discussed above, or
in the cited patent. Item No. 9 refers to the automatic provision for
variant forms, such as "SMITH, JOHN" instead of "JOHN SMITH".

Thus, the invention should not be deemed limited to the particular
embodiments described above. The arrangement of components in Figure 1
is only exemplary; many other combinations of computers, databases, and
modems can be used, only some of which have been described explicitly.
The embodiments of Figures 2, 3 and 4 can be used together or separately,
and, as explained above, some features of each embodiment can be used
alone. All such variations should be deemed within the spirit and scope
of the following claims.

formed substantially simultaneously.

6. A method of supplying information to a user, from a plurality of databases, comprising the steps of accepting, from the user, a search request, the search request including at least one word for which the user desires to search, determining an area of interest relating to the search request, automatically selecting a set of databases, the set being determined by said area of interest, the databases in said set being related to said area of interest, and automatically executing the search request in each member of said set of databases.

7. The method of Claim 6, wherein the step of determining an area of interest comprises selecting an area of interest from at least one menu.

8. The method of Claim 6, wherein the searches are performed sequentially in each of said databases.

9. The method of Claim 6, further comprising the step of displaying, to the user, the results of the searches performed.

10. The method of Claim 6, wherein the searches are performed substantially simultaneously in each of said databases.

11. A system for supplying information to a user from databases, the user being located at a terminal, the system comprising a computer, the computer being connected through a first modem to the user's terminal, the computer being connected through at least one second modem to an outgoing telecommunications line, the computer being programmed to accept, from the user, a search request, the search request comprising at least one word for which the user desires to search, the computer comprising means for determining an area of interest related to the search request, the computer also comprising means for associating at least two databases with said area of interest, the computer also being programmed

37

15. A method of obtaining information from any one of a plurality of databases, each database having a command set for conducting the operation of searches in the database and for displaying the results of such searches, the command sets for at least two databases being different, the method comprising the steps of:

a) specifying a database, a search request and at least one command, the search request including at least one word desired to be searched in the database, the command being selected from a set of standardized commands,

b) translating the command selected from the standardized set into a command which can be recognized by the selected database, and

c) executing the search request in the selected database, according to the translated command.

16. The method of Claim 15, wherein the databases contain information arranged in searchable fields, and wherein the standardized command set includes commands for retrieving information from specified fields, and wherein the translating step includes translating a standardized command relating to information in one of said fields into a command which can be recognized by the selected database.

17. In a system for retrieval of information from a database, the database being located on at least two different database families, the system including a computer, a first modem connected to accept commands from a user located at a terminal, and a second modem connected to an outgoing telecommunications line, the computer being capable of establishing connection with a plurality of database families, using the second modem and the telecommunications line, the computer being programmed to accept a search request from the user, the search request comprising at least one word for which the user desires to search, the computer com-

mation being supplied to a user, the method comprising the steps of:

a)   transmitting a search request to the database, the search request comprising at least one word for which the user desires to search,

b)   determining the number of documents retrieved from the database which fulfill the search request,

c)   determining whether the number of documents retrieved is zero or whether it is greater than a predetermined value, and

d)   displaying to the user, if the number of retrieved documents is zero or greater than said predetermined value, the numbers of documents retrieved by at least one word of the search request, and suggesting to the user that the search can be modified and resubmitted.

22.   The method of Claim 21, wherein step (d) comprises the step of identifying whether any of the words of the search request were not found in any document of the database, and alerting the user if such condition is true.

23.   The method of Claim 21, wherein the suggesting step is selected from the group consisting of suggesting that the user enter a related word instead of a word of the search request, suggesting that the user correct the spelling of a word of the search request, suggesting that the user delete a word from the search request, and suggesting that the user submit a search which seeks documents in which a word of the search request appears in any portion of a document.

24.   In a system for retrieving information from at least one database, the system including a computer which is programmed to accept a search request from a user, the search request including at least one word for which the user desires to search, the computer being programmed to retrieve documents which fulfill the user's search request, the im-

determining the number of documents containing a word of the search re-
quest in a given field of the documents, and ranking the documents in
order of the presumed importance of each field.

28.   The method of Claim 27, wherein the ranking step includes the
step of determining the number of words of the search request which ap-
pear in the same field of a document, and ranking the documents in order
of said number of words.

29.   The method of Claim 26, wherein the ranking step includes the
step of determining the number of documents which contain combinations of
words of the search request, separated by various numbers of words, and
ranking the documents such that the most relevant documents are those in
which the words of the search request appear closest together.

30.   In a system for retrieving information from at least one data-
base, the system including a computer which is programmed to accept a
search request from a user, the search request including at least one
word for which the user desires to search, the computer being programmed
to retrieve documents which fulfill the user's search request, the im-
provement wherein the computer is programmed to rank the retrieved docu-
ments in order of relevance, without examining the documents but only by
determining the number of documents containing various words of the
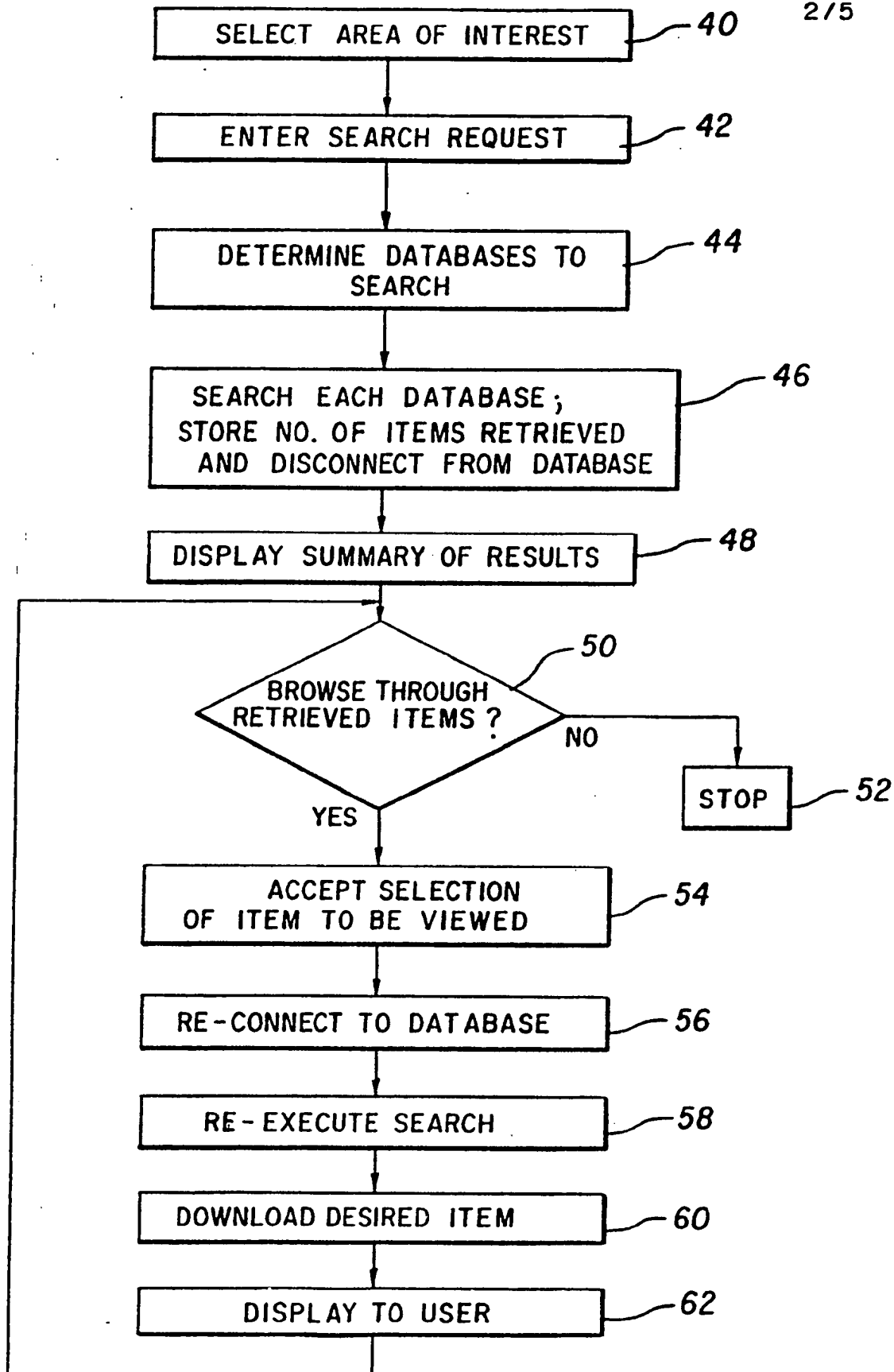search request.

```
┌─────────────────────────────┐
│   SELECT AREA OF INTEREST    │──── 40
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│    ENTER SEARCH REQUEST      │──── 42
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│   DETERMINE DATABASES TO     │──── 44
│          SEARCH              │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────────┐
│   SEARCH EACH DATABASE;          │──── 46
│   STORE NO. OF ITEMS RETRIEVED   │
│   AND DISCONNECT FROM DATABASE   │
└─────────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│  DISPLAY SUMMARY OF RESULTS  │──── 48
└─────────────────────────────┘
              │
              ▼
           ╱──────╲
          ╱  BROWSE ╲  ──── 50
         ╱  THROUGH   ╲        NO    ┌────────┐
        ╲ RETRIEVED   ╱──────────────│  STOP  │──── 52
         ╲  ITEMS ?  ╱               └────────┘
          ╲────────╱
              │ YES
              ▼
┌─────────────────────────────┐
│     ACCEPT SELECTION         │──── 54
│    OF ITEM TO BE VIEWED      │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│    RE-CONNECT TO DATABASE    │──── 56
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│      RE-EXECUTE SEARCH       │──── 58
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│    DOWNLOAD DESIRED ITEM     │──── 60
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│       DISPLAY TO USER        │──── 62
└─────────────────────────────┘
```

FIG. 2

STANDARD COMMAND	MEANING	BRS	EQUIVALENTS DIALOG	VU/TEXT
FIND	SEARCH DATABASE	..SEARCH	SELECT	S
LIST	REVIEW DATABASE DICTIONARY	ROOT	EXPAND	DICT
PHOTO	ORDER PHOTOCOPIES OF ARTICLES	..ORDER	ORDER ITEM	✶
REVIEW	REVIEW SETS	..DISPLAY	DISPLAY SETS	RECALL ALL
SHOW	DISPLAY RESULTS	..PRINT	TYPE	PL
STOP	END SESSION	..OFF	LOGOFF	OFF
TOTAL	SHOW SESSION INFORMATION	..TIME	COST	TOT
COMMANDS	DISPLAY COMMANDS	..C	?EXPLAIN	?
INFO	ADDITIONAL HELP	✶	?(TOPIC)	HELP

✶ NOT KNOWN OR NOT AVAILABLE.

FIG. 4

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US90/00037

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(5): G06F 15/403

U.S. CL.: 364/282.4

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC (5):	G06F 15/21, 15/40, 15/403, 15/409.
U.S. CL.:	364/200 364/282.4, 364/900, 364/974

Documentation Searched other than Minimum Documentation
to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category [*]	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
Y	Searching DIALOG The Complete Guide, "Chapter 9 Searching Multiple Files" January 1988-- See pp. (9-1), (9-4), (9-6), (9-14), (9-15) (9-24), (9-25)	1,2,5,6,8-11 13,18,22
Y	Searching DIALOG The Complete Guide, "chapter 1 Introduction to Dialog" August 1987 See pages (1-2) and (1-5)	3,4,20
Y	US, A, 4,774,655 Kollin et al. 27 September 1988 (27.09.88) See figures 1,3a,3b, abstract, Col. 2 (line 45-et seq.), Col, 4 (line 47-et seq.)	7,11,12-17 19
Y,P	US, A, 4,829,423 Tennant et al. 09 May 1989 (09.05.89) See figures 1-16, Col. 3 (line 14-et seq.)	21,23-26
Y	Searching DIALOG The Complete Guide "chapter 8 DIALOG Commands" See pages (8-Sort 1) to (8-Sort6)	26-30

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search

20 MARCH 1990 (20.03.1990)

Date of Mailing of this International Search Report

27 APR 1990

International Searching Authority

ISA/US

Signature of Authorized Officer

ROBERT B. HARRELL

Form PCT/ISA/210 (second sheet) (Rev.11-87)